A Recurrent Neural Network Approach in Predicting Daily Stock Prices

An Application to the Sri Lankan Stock Market

A.J.P. Samarawickrama
Department of Computer Science
Faculty of Applied Sciences
University of Sri Jayewardenepura
Nugegoda, Sri Lanka
jayanath@sjp.ac.lk

T.G.I. Fernando
Department of Computer Science
Faculty of Applied Sciences
University of Sri Jayewardenepura
Nugegoda, Sri Lanka
gishantha@dscs.sjp.ac.lk

***Abstract*— Recurrent Neural Networks (RNNs) is a sub type of neural networks that use feedback connections. Several types of RNN models are used in predicting financial time series. This study was conducted to develop models to predict daily stock prices of selected listed companies of Colombo Stock Exchange (CSE) based on Recurrent Neural Network (RNN) Approach and to measure the accuracy of the models developed and identify the shortcomings of the models if present. Feedforward, Simple Recurrent Neural Network (SRNN), Gated Recurrent Unit (GRU) and Long Short Term Memory (LSTM) architectures were employed in building models. Closing, High and Low prices of past two days were selected as input variables for each company. Feedforward networks produce the highest and lowest forecasting errors. The forecasting accuracy of the best feedforward networks is approximately 99%. SRNN and LSTM networks generally produce lower errors compared with feedforward networks but in some occasions, the error is higher than feed forward networks. Compared to other two networks, GRU networks are having comparatively higher forecasting errors.**

Keywords*— *Recurrent Neural Networks, LSTM, SRNN, GRU, Colombo Stock Exchange

I. Introduction

In today's world, financial markets have a significant impact on the economies in which they operate. In most of the modern economies business organizations heavily rely on the funds generated by these financial markets. Therefore, analyzing the behavior and performance of financial markets has become a major research field. These analyses include predicting prices of securities (E.g. Stocks, bonds, etc.), foreign exchange rates, market indicators, returns of securities and trading volumes, classification of stock, etc [1]. In the past (before 1990s) traditional statistical models like linear regression, traditional time series forecasting, technical analysis and fundamental analysis were employed for financial market analyses and predictions [2].

Due to the uncertainty, complexity and non-linearity of the financial market data traditional statistical methods were not successful enough to do analyses and predictions based on financial market data [3]. So, the focus has been shifted from linear models to non linear models including machine learning techniques like Artificial Neural Networks (ANN), Support Vector Machines (SVM), Genetic Algorithms (GA), etc. [4].

Artificial Neural Networks (ANNs) is a branch of Artificial Intelligence (AI). ANNs consist of computational models and electronic circuits that simulate the functions of human central nervous system [5]. The beginning of ANNs was the invention of a simple neural network model with electrical circuits by McCulloch and Pits in 1943 [6]. Today ANNs are applied to many areas like time series predictions, classification and pattern recognition, medical diagnoses and hardware device controlling, etc. [7]. Modeling financial market data (especially stock price prediction) is one of the major areas that ANNs are applied [8].

Recurrent Neural Networks (RNN) is a sub type of neural networks that use feedback connections. Several types of RNN models are employed in predicting financial time series. This study was conducted in order to develop models to predict daily stock prices of selected listed companies of Colombo Stock Exchange (CSE) based on Recurrent Neural Network (RNN) Approach and to measure the accuracy of the models developed and identify the shortcomings of the models if present.

Section II of this paper presents a review of research work related with this study. Section III describes the theoretical framework and section the research methodology. Section IV discusses the results and section V concludes the paper.

II. Literature Review

A study was done in 1988 to predict IBM daily stock returns [9]. In 1989 it was found that Neural Networks (NN) are powerful and ideal for forecasting and classifying financial market data as they have the capability in recognizing patterns in complex, chaotic data sets and dealing with non-linearity [10].

978-1-5386-1676-5/17/$31.00 ©2017 IEEE

Number of studies has been carried out to analyze and predict financial market data using NNs. There are several neural network architectures like Multi-Layer Perceptrons (MLPs), Recurrent Neural Networks (RNNs), Probabilistic Neural Networks, Neuro-Fuzzy Systems, etc. that employ several algorithms like backpropagation, quick propagation, deep learning, Levenberg-Marquardt (LM) algorithm. Multi Layer Perceptron (Feedforward Neural Network) is the simplest and widely used architecture used with backpropagation algorithm which is the simplest algorithm used in ANNs [11].

Most of the studies focused on stock market analysis have employed the feedforward network with backpropagation algorithm but it has several limitations as it is too slow for practical problems. Due to these limitations several other neural network structures/algorithms like Recurrent Neural Networks, Deep Neural Networks, Radial Basis Functions, etc. are used in stock market analyses [11].

Recurrent Neural Networks (RNN) use the backpropagation in learning but it has a feedback mechanism in its nodes. Because of that RNN models can predict stock prices based on more recent history [12].

A study was conducted in 2010 using Multi-Layer Perceptron (MLP) and Elman neural network to predict the stock values of 1094 companies traded in the Tehran Stock Market (during 2000 to 2005) based on stock share value history and found that MLP is more accurate in stock value prediction while Elman network is more accurate in predicting directional changes in stock values [13].

In 2012, a model was developed based on RNN to predict stock prices of S&P 500 index in Indian stock market using Echo State Networks (ESN). Data was sourced from Yahoo Finance. Daily stock prices from late 2004 to early 2009 were collected for the model building. Current and 5 day history of the stock price, 5-, 10-, 15- and 20-day moving averages, volume and the S&P500 index were the variables of the study. The model had the ability to predict stock prices with higher frequency fluctuations [14].

A study was conducted in 2011 to predict closing share prices of two listed companies in Colombo Stock Exchange (CSE) using ANN trained with Levenberg-Marquardt algorithm and Support Vector Regression (SVR). Closing share prices from January, 1999 to July, 2010 were selected as data for this study. In this study NN models showed high performance for large data sets and SVR models showed high performance for small data sets [4].

III. THEORETICAL FRAMEWORK

A. Neural Network Topologies

Mainly, there are two NN topologies namely feedforward and feedback neural networks. Feedforward Neural Network is the most commonly used topology where the data flow from input layer to output layer is strictly feedforward (Fig. 1). Either the network has multiple layers or not there are no any feedback connections [15].

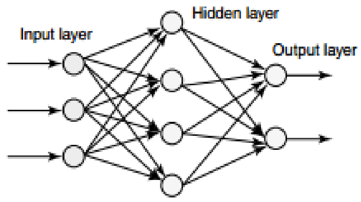

Figure 1: Feedforward Neural Network **[20]**

B. Recurrent Neural Networks

In 1983, Hopfield introduced feedback connections to the feedforward backpropagation networks (Fig. 2). This was the introduction of Recurrent Neural Networks (RNNs). As Recurrent Neural Networks are having at least single feedback connection they have the ability to store correct pattern and they were based on more recent history than the past [16].

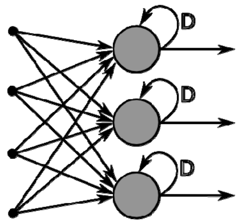

Figure 2: Recurrent Neural Network **[21]**

There are several Recurrent Neural Network (RNN) architectures like Fully Recurrent Neural Networks, Simple Recurrent Neural Networks (Elman Networks and Jordan Networks), Hopfield Networks, Long Short Term Memory (LSTM) Networks, Echo State Networks, Continuous-time RNN and Bi-directional RNN, etc.

C. Simple Recurrent Neural Networks

Jeff Elman developed a special type of RNN architecture (Fig. 3). In this architecture a three layer network with the addition of a set of "context units" is used. There are connections from hidden layer to the neurons in the context layer with fixed weights. This type of network (called as Elman Network) is better to perform sequence predictions than feedforward networks [17]. Jordan Network architecture is developed by Michael I. Jordan. It is similar to Elman networks but the context units are fed from the output layer instead of the hidden layer. Elman and Jordan networks both are known as Simple Recurrent Neural Networks (SRNN) [17].

Although Simple RNNs are capable of learning sequential time series patterns, they have vanishing gradient problem that affects their performance [17].

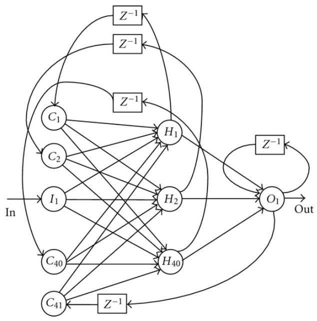

Figure 3: Simple Recurrent Neural Networks **[23]**

D. *Long Short Term Memory*

As a solution to the “vanishing gradient problem,” German researchers Sepp Hochreiter and Juergen Schmidhuber proposed the Long Short Term Memory (LSTM) units in the mid-90s (Fig. 4). LSTM network is a variation of RNNs that help preserve the error that can be propagated through time and layers. An LSTM network contains LSTM blocks instead of/ in addition to regular network units (neurons). An LSTM block contains 3 gates namely input gate, output gate and forget gate [18].

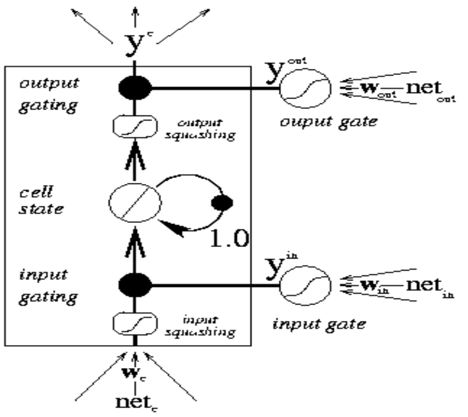

Figure 4: Architecture of an LSTM Unit **[18]**

E. *Gated Recurrent Units*

Gated Recurrent Units (GRU) is another type of RNNs, hypothesized in 2014 (Fig. 5). GRU is a gating mechanism in recurrent neural networks. A GRU is basically an LSTM but the difference is: a GRU hasn’t an output gate [19].

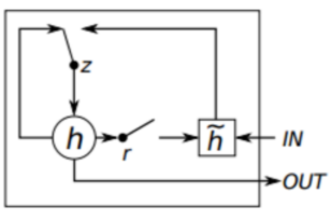

Figure 5: Architecture of a Gated Recurrent Unit **[19]**

IV. METHODOLOGY

A. *Selection of Data*

From 297 listed companies of Colombo Stock Exchange (CSE), 3 companies were selected for this study. These companies were selected from 3 industries/sectors with the highest annual share trading volume from 20 sectors of CSE. Companies were selected based on the stability of the stock price and the performance in the market (trading volume and frequency). Commercial Bank Plc. (COM) (Banks, Finance & Insurance Sector), Royal Ceramics Limited (RCL) (Manufacturing Sector) and John Keels Holdings (JKH) (Diversified Holding Sector) were selected for this study. Data from these companies during the period from 2002/01/01 to 2013/06/30 were selected for model building.

B. *Selection of Variables*

For model building Closing, High and Low prices of past two days were selected as input variables for each company as they were significantly correlated with the output variable Closing Price.

Accordingly the input variables are as follows (t – current time):

C(t-1) – the closing price of the day t-1
C(t-2) – the closing price of the day t-2
L(t-1) – the low price of the day t-1
L(t-2) – the low price of the day t-2
H(t-1) – the high price of the day t-1
H(t-2) – the high price of the day t-2

Table I: Input and Output Variables

Input Variables	C(t-1), C(t-2)
	L(t-1), L(t-2)
	H(t-1), H(t-2)
Output Variable	C(t)

C. *Neural Network Topology*

For this study, 3 popular neural network architectures were selected for model building. Those were Simple Recurrent Neural Network (SRNN), Long Short Term Memory (LSTM) and Gated Recurrent Unit (GRU). In addition to these neural network architectures Feed Forward Neural Network (Multi-Layer Perceptron – MLP) architecture was also employed for comparison. For each of neural network architectures, 10 neural network models were developed (by varying the number of hidden units from 2 to 11). So, there were 40 neural network models for each company. Total number of neural network models developed for this study was 120.

Each model has an input layer with 6 input neurons, a hidden layer with 2 to 11 hidden neurons and an output layer with 1 output neuron. The following table describes the internal architecture of each of 10 models developed using a single architecture for a single company.

Table II: Number of Neurons in Each Model

Model Number	No. of Input Neurons	No. of Hidden Neurons	No. of Output Neurons
1	6	2	1
2	6	3	1
3	6	4	1
4	6	5	1
5	6	6	1
6	6	7	1
7	6	8	1
8	6	9	1
9	6	10	1
10	6	11	1

D. *Data Preprocessing*

In this study, the input and output data was normalized into the range of [0, 1] using the following formula.

$$X_p = \left(\frac{X - \min X}{\max X - \min X} \right) \tag{1}$$

X_p – Preprocessed data value
X – Unprocessed data value
$minX$ – Minimum value
$maxX$ – Maximum value

E. *Training and Performance Evaluation*

Most recent data were selected for testing, next recent dataset was selected for the validation and the rest were selected for the training. The size of the test set was 15% of the whole dataset and the size of the validation set is also 15% of the whole dataset. The rest of the dataset was used to train the models. The following table shows the sizes of training, validation and testing sets of each company.

Table III: Sizes of Training, Validation and Testing Data Sets

Data Set	JKH		RCL		COM	
	Size	%	Size	%	Size	%
Training	1915	69%	1884	70%	1879	70%
Validation	420	15%	413	15%	412	15%
Testing	412	15%	405	15%	404	15%
Total	2747	99%	2702	100%	2695	100%

Keras deep learning library was used to build and train neural networks on windows platform. The following parameters and activation functions were used to train neural networks.

Table IV: Training Parameters for Neural Networks

Parameter	MLP	SRNN	LSTM	GRU
Initialization function	uniform	glorot uniform	glorot uniform	glorot uniform
Inner Initialization function	-	orthogonal	orthogonal	orthogonal
Activation function	relu	softsign	softsign	softsign
Inner activation function	-	hard sigmoid	hard sigmoid	hard sigmoid
Error (loss) function	MSE	MSE	MSE	MSE
Optimizer/Training algorithm	adam	rmsprop	rmsprop	rmsprop

Each of the 40 neural network models was trained for 5000 iterations. So, the stopping condition for all the networks was 5000 iterations.

After training, predictions were done using the test dataset. To measure the performance of each neural network model, Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE) were calculated in order to measure the accuracy of each model. The model with lowest MAD/MAPE was considered as the best model.

$$MAD = \frac{1}{|ValidationSet|} \sum_{for.all.days \in ValidationSet} \left| price_{forecast}^{tomorrow} - price_{real}^{tomorrow} \right| \tag{2}$$

$$MAPE = \frac{1}{|ValidationSet|} \sum_{for.all.days \in ValidationSet} \left| \frac{price_{forecast}^{tomorrow} - price_{real}^{tomorrow}}{price_{real}^{tomorrow}} \right| \tag{3}$$

To compare predicted prices with actual prices, the values resulted from testing were converted again to the raw format.

V. RESULTS OF THE STUDY

Table V: Best Networks of Each Company According to Forecast Error

NN Architecture	JKH	RCL	COM
MLP	3	3	6
SRNN	9	7	7
GRU	8	8	4
LSTM	7	4	9

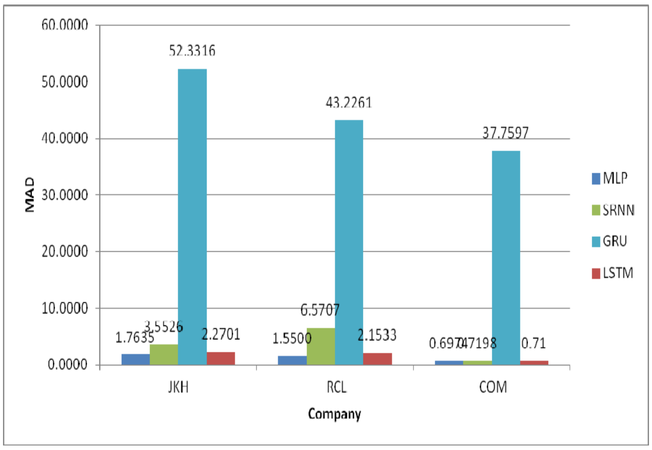

Figure 6: Forecast Errors (MAD) - Best Models of Each Company

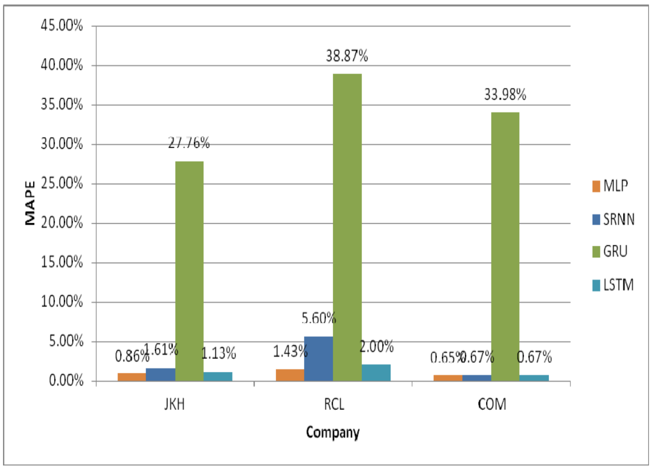

Figure 7: Forecast Errors (MAPE) - Best Models of Each Company

According to the tables and figures in the section 5, MLP models (in each company) produce the lowest forecasting errors. From the 40 models of the John Keels Holdings (JKH) MLP Model 2 (with 3 hidden units) has the lowest forecasting error (MAD – 1.7636; MAPE – 0.86%). From the 40 models of the Royal Ceramics Limited (RCL) MLP Model 2 (with 3 hidden units) has the lowest forecasting error (MAD – 1.5500; MAPE – 1.43%). From the 40 models of the Commercial Bank Plc (COM) MLP Model 5 (with 6 hidden units) has the lowest forecasting error (MAD – 0.6974; MAPE – 0.65%).

From the 30 recurrent neural network models of the John Keels Holdings (JKH), LSTM Model 2 (with 7 hidden units) has the lowest forecasting error (MAD – 2.2701; MAPE – 1.13%). From the 30 recurrent neural network models of the Royal Ceramics Limited (RCL), LSTM Model 2 (with 4 hidden units) has the lowest forecasting error (MAD – 2.1533; MAPE – 2.00%). From the 30 recurrent neural network models of the Commercial Bank Plc (COM), LSTM Model 5 (with 9 hidden units) has the lowest forecasting error (MAD – 0.7100; MAPE – 0.67%).

The best Simple Recurrent Neural Network (SRNN) of each company produces MAD ranging from 0.72 – 6.57 (MAPE ranging from 0.67% - 5.60%). Gated Recurrent Units (GRU) produce relatively higher errors.

From all the 120 neural network models the MLP Model 5 of the Commercial Bank is the model having the lowest forecasting error. From all the 80 recurrent neural network models the LSTM Model 8 of the Commercial Bank is the model having the lowest forecasting error. So, it can be concluded that among all the NN models developed MLP Model 5 of the Commercial Bank is the best and among all the RNN models developed LSTM Model 8 of the Commercial Bank is the best model.

VI. Discussion, Conclusion and Future Research Work

The previous studies discussed in Literature Review (Section II) have employed several fundamental, technical and other types of variables for model building. But in this study only few technical variables were considered as inputs. Above mentioned previous studies have employed MLP with backpropagation, MLP with Levenberg-Marquardt algorithm, Elman Neural Networks and Echo State Networks. In this study MLP, Simple RNN, LSTM and GRU architectures have been employed. In previous studies number of input variables and number of neurons in hidden layers were varied to find the best model. In this study only the number of hidden neurons was changed while fixing the number of input variables to find the best model.

In this study, when considering the forecast error (or the test error) MLP models produce the highest and the lowest errors. The forecasting accuracy of the best feedforward networks is approximately 99%. SRNN and LSTM networks generally produce lower errors compared with feedforward networks but in some occasions, the error is higher than feedforward networks. Compared to other two networks, GRU networks are producing comparatively higher forecasting errors.

When considering the results of previous studies in most of the studies recurrent neural network models (especially LSTM) produce the best results. But in this study MLP models produce the best results. This is because; in this study data of only past two days were selected for inputs. If the number of past days considered for input variable selection was increased, RNN models would produce best results.

This study is only limited to three companies in manufacturing, diversified and banking sectors. But model development can be extended to other sectors as well as other companies in these sectors. And there are varieties of technical, fundamental and other types of variables that can be used in developing models. This study predicts only the closing prices but models can be developed to predict open, low, high prices, transaction volumes and return on securities, etc.

References

[1] Y. Singh and A. S. Chauhan, "Neural Networks in Data Mining," *Journal of Theoretical and Applied Information Technology,* pp. 37-42, 2009.

[2] R. Lawrence, "Using Neural Networks to Forecast Stock Market Prices," 1997.

[3] B. Malkiel, A Random Walk Down Wall Street, New York: Norton &

Company, 1985.

[4] D. Rangana, N. Chandrasekara and C. Tilakaratne, "Comparison of Support Vector Regression and Artificial Neural Network Models to Forecast Daily Colombo Stock Exchange," 2011.

[5] C. Janssen, "What is Aritificial Neural Network," [Online]. Available: http://www.techopedia.com/. [Accessed 17 May 2015].

[6] W. S. McCulloch and W. H. Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity," *Bulletin of Mathematical Biophysics,* vol. IV, pp. 115-133, 1943.

[7] S. B. Maind and P. Wankar, "Research Paper on Basic of Artificial Neural Network," *International Journal on Recent and Innovation Trends in Computing and Communication,* vol. II, no. 1, pp. 96-100, 2014.

[8] M. Zekic, "Neural Network Applications in Stock Market Predictions," in *9th International Conference on Information and Intelligent Systems*, 1998.

[9] H. White, "Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Returns," in *the Second Annual IEEE Conference on Neural Networks*, New York, 1988.

[10] G. Cybenko, "Approximation by Superpositions of a Sigmodial Function," *Mathematics of Control, Signals and Systems,* vol. II, pp. 303-314, 1989.

[11] J. B. Singh, "Current Approaches in Neural Network Modelling of Financial Time Series," Bangalore, 2009.

[12] J. Roman and A. Jameel, "Backpropagation and Recurrent Neural Networks in Financial Analysis of Multiple Stock Market Returns," in *29th Annual Hawaaii International Conference on System Sciences*, 1996.

[13] M. P. Naeini, H. Taremian and H. B. Hashemi, "Stock Market Value Prediction Using Neural Networks," in *International Conference on Computer Information Systems and Industrial Management Applications (CISIM)*, 2010.

[14] A. Bernal, S. Fok and R. Pidaparthi, "Financial Market Time Series Prediction with Recurrent Neural Networks," 14 December 2012.

[15] L. Jacobson, "Introduction to Artificial Neural Networks Part 2 - Learning," 26 March 2014. [Online]. Available: http://www.theprojectspot.com/tutorial-post/introduction-to-artificial-neural-networks-part-2. [Accessed 10 February 2016].

[16] K. D. Gunawardana, An Introduction to Artificial Neural Networks for Accounting and Finance Modelling, Colombo: Author Publication, 2009.

[17] "Recurrent neural network," 21 February 2016. [Online]. Available: https://en.wikipedia.org/wiki/Recurrent_neural_network. [Accessed 28 February 2016].

[18] "Long short-term memory," 9 July 2016. [Online]. Available: https://en.wikipedia.org/wiki/Long_short-term_memory. [Accessed 27 August 2016].

[19] "Gated recurrent unit," 12 July 2016. [Online]. Available: https://en.wikipedia.org/wiki/Gated_recurrent_unit.

[20] Cheshire Engineering Corporation, "Basic Concepts for Neural Networks," [Online]. Available: http://www.cheshireeng.com/Neuralyst/nnbg.htm. [Accessed 31 January 2016].

[21] "Recurrent Neural Networks Tutorial, Part 1 – Introduction to RNNs," 17 September 2015. [Online]. Available: http://www.wildml.com/. [Accessed 22 January 2016].

[22] B. A. Pearlmutter, "Dynamic Recurrent Neural Networks," December 1990.

[23] J. Szkoła, K. Pancerz and J. Warchoł, "Recurrent Neural Networks in Computer-Based Clinical Decision Support for Laryngopathies: An Experimental Study," Computational Intelligence and Neuroscience, 2011.